

Issues in XAI #5: Understanding Black Boxes

— Interdisciplinary Perspectives

TU Dortmund — September 5–7, 2022

Schedule

Sunday, September 4

19.00 Dinner: No More Rice

Märkische Str. 62, 44141 Dortmund; <https://www.nomorerice.de/>

Monday, September 5

11.00 – 11.10: Welcome & Introduction

11.10 – 12.20: Andrés Páez (philosophy): “Idealization and Non-Factive Understanding in Machine Learning” (chair: Chiara Balestra)

— 15 minutes coffee break —

12.35 – 13.45: Nicole Krämer (psychology): “Understanding versus trust: What do users need when interacting with AI systems?” (chair: Andrés Páez)

— 80 minutes lunch break —

15.05 – 16.15: Sara Mann (philosophy): “Understanding via Exemplification in XAI” (chair: Nicole Krämer)

— 15 minutes coffee break —

16.30 – 17.40: Chiara Balestra (computer science): “Coalitional Game Theory to explain the structure of the data” (chair: Philipp Cimiano)

19.30 Dinner: Hövels

Hoher Wall 5-7, 44137 Dortmund; <https://www.hoevels-hausbrauerei.de/>

Tuesday, September 6

09.30 – 10.40: Emanuele Ratti (philosophy): “An Integrative and coherentist approach to XAI in scientific research: models ‘for’ and systems of practice” (chair: Andreas Sesting-Wagenpfeil)

— 15 minutes coffee break —

10.55 – 12.05: Tim Hunsicker (psychology): “Unexplainable Accuracy or Explainable Faultiness? Investigating the Accuracy-Transparency Trade-Off of AI-Based Systems” (chair: Emanuele Ratti)

— 5 minutes coffee break —

12.10 – 13.20: Florian J. Boge (philosophy): “Functional Concept Proxies and the Actually Smart Hans Problem: What’s Special About Deep Neural Networks in Science” (chair: Tim Hunsicker)

— 80 minutes lunch break —

14.40 – 15.50: Mieke Boon (philosophy): “Where to Locate the Explainability of Explainable Machine Learning?” (chair: Florian J. Boge)

— 15 minutes coffee break —

16.05 – 17.15: Sabine Ammon (philosophy): “When is a Black-Box a Problem? Epistemic Coercion and Epistemic Sovereignty in Explainable Artificial Intelligence (XAI)” (chair: Mieke Boon)

— 5 minutes coffee break —

17.20 – 18.30: Andreas Sesting-Wagenpfeil (law): “Discrimination and the AI Act” (chair: Sabine Ammon)

20.00 Dinner: 60 Seconds to Napoli

Markt 5, 44137 Dortmund; <https://60secondstonapoli.de/dortmund/>

Wednesday, September 7

10.00 – 11.10: Philipp Cimiano (computer science): “Counterfactual explanations for image classification tasks with an application in medical decision making” (chair: Juan M. Durán)

— 15 minutes coffee break —

11.25 – 12.35: Nadine Schlicker (psychology): “Men Are From Mars, Machines Are From Venus — Explaining Audio Processing of Deep Neural Networks” (chair: Sara Mann)

— 15 minutes coffee break —

12.50 – 14.00: Juan M. Durán (philosophy & computer science): “Thinking Outside of the (Black) Box: Computational Reliabilism and Epistemic Trust” (chair: Nadine Schlicker)